OPTIMAL CONTROL OF A BROWIAN STORAGE SYSTEM

by

J. MICHAEL HARRISON   and   ALLISON J. TAYLOR

TECHNICAL REPORT NO. 69

September 1, 1976

DEPARTMENT OF OPERATIONS RESEARCH

STANFORD   UNIVERSITY

STANFORD, CALIFORNIA

OPTIMAL CONTROL OF A BROWNIAN STORAGE SYSTEM

J. Michael Harrison

Allison J. Taylor

Stanford University

McMaster University

1. Introduction and Summary

We consider in this paper two highly structured problems of optimal stochastic control. The two problems will be precisely formulated, and optimal control policies of a simple form will be explicitly computed, in Sections 3 and 4. In this section we present an informal description of each problem and its solution, suppressing all technical detail.

Let $X = \{X(t), t \geq 0\}$ be a Brownian Motion with starting state $x \geq 0$, drift μ and variance $\sigma^2 > 0$. Thus $E[X(t)] = x + \mu t$ and $Var[X(t)] = \sigma^2 t$. We define a control to be a non-decreasing process $Y = \{Y(t), t \geq 0\}$, with $Y(0) \geq 0$, which is a non-anticipating functional (of X). Thus, for each $t \geq 0$, the partial control history $\{Y(u), 0 \leq u \leq t\}$ may depend on $\{X(u), 0 \leq u \leq t\}$ and possibly on other information as well, but it may not depend on $\{X(t+u) - X(t), u \geq 0\}$. (See Section 2 for a precise definition.) In each of our problems, the objective is to find an input control Y and an output control Z which maximize expected discounted reward (over an infinite planning horizon) subject to the constraint that $W(t) = X(t) + Y(t) - Z(t) \geq 0$ for all $t \geq 0$ (almost surely). It is the hypothesized structure of costs and rewards that differs in the two problems.

Given a constant $k > 1$, the objective in our first problem is to find an admissible control policy (Y,Z) which maximizes

$$E[Z(0) + \int_0^\infty e^{-\alpha t} \, dZ(t)] - kE[Y(0) + \int_0^\infty e^{-\alpha t} \, dY(t)]$$

where $\alpha > 0$ is the interest rate. By way of interpretation, we imagine a storage system (such as an inventory or a bank account) whose content evolves as the Brownian Motion X in the absence of any control. In particular, $X(0)$ represents the initial content of the system. The controller may at any time withdraw material from the system, and $Z(t)$ represents the total withdrawal during the interval $[0,t]$, or cumulative output up to time t. He receives a reward of one dollar for each unit of material withdrawn. If the content of the system falls to zero, however, then the controller is obliged to inject material into the system so as to keep the net content positive, and he incurs a cost of $k > 1$ dollars for each unit of material injected. We interpret $Y(t)$ as the total injection during the interval $[0,t]$, or the cumulative input up to time t. We call $W = X + Y - Z$ the controlled process and X the uncontrolled process.

In Section 3 it will be shown that an optimal policy for this first problem is the minimal pair of controls (Y,Z) which achieves $0 \le W(t) \le S$ for all $t \ge 0$ (almost surely), where S is the unique positive solution of a certain transcendental equation. Thus the controller should withdraw only as much material as is required to keep the net content below S, and he should inject the minimum amount necessary to keep the net content positive. The optimal controls are

explicitly described in terms of certain maximum and minimum functionals

applied to X, and the corresponding controlled process W behaves as

the Brownian Motion X modified by (instantaneously) reflecting barriers

at zero and S. The optimal controls are (almost surely) continuous but

not absolutely continuous, due to the unbounded variation of Brownian

paths. Thus one cannot describe the optimal policy in terms of input and

output rates. Still it is in a certain sense a bang-bang policy, as

one would expect with our linear cost structure.

Our second problem differs only in that we require the input control

Y to be a (random) step function, and we assume that a fixed change of

K > 0 dollars is incurred each time an input jump occurs. (This is

additional to the proportional charge of k dollars per unit of input.)

In Section 4 it will be shown that the optimal policy is as follows. Each

time that the net content W hits zero, the controller increases the

cumulative input Y by s units, where s is the unique positive solu-

tion of a certain transcendental equation. Between these input events,

the controller increases the cumulative output Z by the minimum amounts

necessary to keep the controlled process below level s+S, where the

positive constant S is the same as in our first problem.

Section 2 contains a number of formal definitions and some

important preliminary propositions. The central results are stated and

proved in Sections 3 and 4. In Section 5 we discuss the application of

our results to problems of inventory control and stochastic cash manage-

ment. In particular, it is shown that our formulation need be altered

only trivially to include problems where the controller continuously

incurs holding costs at a rate proportional to the net content of the

system $W(t)$. We also discuss the relationship between our model(s)
and other (approximate) diffusion formulations that have been suggested
for such problems. In Section 6 we discuss the difference between our
formulation and various other theories of optimal stochastic control.

2. Preliminaries

Let $\mathbb{R} = (-\infty, \infty)$, $\mathbb{R}^+ = [0, \infty)$, $I = \{0, 1, \ldots\}$ and $I^+ = \{1, 2, \ldots\}$.
Throughout the paper, let $\alpha > 0$, $\mu \in \mathbb{R}$, $\sigma^2 > 0$, $k > 1$ and $K > 0$ be
fixed constants. We assume a measurable space $(\Omega, \mathfrak{F})$ on which is
defined a family of probability measures $\{P_x, x \in \mathbb{R}\}$ and a process
$X = \{X(t), t \geq 0\}$ such that X is Brownian Motion with drift μ,
variance $\sigma^2 > 0$, and starting state x with respect to P_x $(x \in \mathbb{R})$.
We denote by E_x the expectation operator associated with P_x. The
following proposition follows easily from standard properties of
Brownian Motion.

<u>Proposition 1</u>. $E_x \sup\{e^{-\alpha t} X(t) : t \geq 0\} < \infty$, $x \in \mathbb{R}$.

We further assume the existence of an increasing family of sub-σ-
algebras $\{\mathfrak{F}_t, t \geq 0\}$ such that X is adapted to $\{\mathfrak{F}_t\}$ and
$X(t+u) - X(t)$ is independent (with respect to P_x for all $x \in \mathbb{R}$)
of $\mathfrak{F}_t$ for all $t \geq 0$ and $u \geq 0$. We say that a random variable T
is a stopping time if $P_x\{0 \leq T < \infty\} = 1$ for all $x \in \mathbb{R}$ and
$\{T \leq t\} \in \mathfrak{F}_t$ for all $t \geq 0$. Using the separability of X and our

assumptions on $\{\mathfrak{F}_t\}$, one easily obtains the following by a standard type of argument.

Proposition 2. If T is a stopping time and Z is measurable $\mathfrak{F}_T$ then, for all $x \in \mathbb{R}$,

$$P_x\{Z + X(T+t) \le y \mid \mathfrak{F}_T\} = P_{Z+X(T)}\{X(t) \le y\}, \qquad t \ge 0, \; y \in \mathbb{R} .$$

Proposition 3. If T is a stopping time, then

$$\alpha E_x \int_0^T e^{-\alpha t} X(t)\, dt + E_x\, e^{-\alpha T}\, X(T) = x + \mu[1 - A(x)]/\mu , \qquad x \in \mathbb{R} ,$$

where $A(x) = E_x\, e^{-\alpha T}$, $x \in \mathbb{R}$.

Proof: Let $f(x) = \alpha E_x \int_0^\infty e^{-\alpha t} X(t)\,dt$, $x \in \mathbb{R}$. Then

$$f(x) = \alpha \int_0^\infty e^{-\alpha t} E_x\, X(t) = \alpha \int_0^\infty e^{-\alpha t}(x+\mu t) = x + \mu/\alpha$$

by Fubini's Theorem. Using this and Proposition 2, we have

$$x + \mu/\alpha - E_x \int_0^T e^{-\alpha t} X(t)\,dt = f(x) - E_x \int_0^T e^{-\alpha t} X(t)\,dt$$

$$= E_x \int_T^\infty e^{-\alpha t} X(t)\, dt = E_x\{e^{-\alpha T} E_x[\int_0^\infty e^{-\alpha t} X(T+t)\,dt \mid \mathfrak{F}_T]\}$$

$$= E_x\, e^{-\alpha T}\, f(X(T)) = E_x\, e^{-\alpha T}[X(T) + \mu/\alpha] ,$$

which completes the proof.

5

If   f   is a real-valued function on some interval subset of $\mathbb{R}$, we define $\mathcal{L}f(x) = \mu f'(x) + \frac{1}{2} \sigma^2 f''(x)$ for all   x   such that $f'(x)$ and $f''(x)$ exist.

Proposition 4.   Let $f : \mathbb{R} \to \mathbb{R}$ be non-decreasing and twice continuously differentiable with $\mathcal{L}f - \alpha f \leq 0$.   Assume $x \in \mathbb{R}$ and let T_* and T^* be stopping times with $P_x\{T_* < T^*\} = 1$.   Finally, let   Z   be measurable $\mathfrak{F}_{T_*}$. Then

$$(1) \qquad E_x\, e^{-\alpha T^*}\, f(Z + X(T^*)) \leq E_x\, e^{-\alpha T_*}\, f(Z + X(T_*))\quad ,$$

if both expectations exist and are finite.

Proof: Assume first that   f   and its first two derivatives are bounded. We begin by proving a slight generalization of the discounted form of Dynkin's identity, c.f. Breiman (1968), equation (16.61).   Let $g(x) = \mathcal{L}f(x) - \alpha f(x)$ for $x \in \mathbb{R}$.   Then a standard result for Markov processes (in this case the Brownian Motion   X   with generator $\mathcal{L}$)   and their resolvents gives us

$$(2) \qquad f(x) = E_x \int_0^\infty e^{-\alpha t}\, g(X(t))\, dt\ , \qquad\qquad x \in \mathbb{R}\ ,$$

cf. Breiman (1968), Theorem 15.51.   Since   Z   is measurable $\mathfrak{F}_{T_*}$, we can combine (2) with Proposition 2 to obtain

$$(3) \qquad E_X \int_{T_*}^{\infty} e^{-\alpha t} \, g(Z + X(t)) dt = E_x \, e^{-\alpha T_*} \int_0^{\infty} e^{-\alpha t} \, g(Z + X(T_*+t)) dt$$

$$= E_x \{ e^{-\alpha T_*} \, E_x [\int_0^{\infty} e^{-\alpha t} \, g(Z + X(T_*+t)) dt \, | \, \mathfrak{F}_{T_*}] \}$$

$$= E_x \, e^{-\alpha T_*} \, E_{Z+X(T_*)} \int_0^{\infty} e^{-\alpha t} \, g(X(t)) dt$$

$$= E_x \, e^{-\alpha T_*} \, f(Z + X(T_*)) \quad .$$

But Z is measurable $\mathfrak{F}_{T^*}$ also, so an identical argument gives

$$(4) \qquad E_x \int_{T^*}^{\infty} e^{-\alpha t} \, g(Z + X(t)) = E_x \, e^{-\alpha T^*} \, f(Z + X(T^*)) \quad .$$

Subtracting (3) from (4) gives

$$E_x [e^{-\alpha T^*} \, f(Z + X(T^*)) - e^{-\alpha T_*} \, f(Z + X(T_*))]$$

$$= E_x \int_{T_*}^{T^*} e^{-\alpha t} \, g(Z + X(t)) \, dt \quad .$$

The right side is non-positive, since $g(x) \leq 0$ for all $x \in \mathbb{R}$, so the proposition is proved.

If either f or one of its derivatives is unbounded, we can easily construct a sequence of bounded functions f_n having two bounded continuous derivatives and such that $f_n(x) = f(x)$ if $|f(x)| < n$ and $|f_n| < |f|$, $n \in I^+$. (The construction is particularly easy with our assumption that f is monotone.) We have shown that (1) holds with f_n in place of f, and obviously $f_n \to f$, so (1) holds by dominated convergence if both expectations exist and are finite. This completes the proof.

We define a non-anticipating functional to be a process $Z = \{Z(t), t \geq 0\}$ on $(\Omega, \mathfrak{F})$ taking values in $D[0,\infty)$ and such that $Z(t)$ is measurable $\mathfrak{F}_t$ for each $t \geq 0$. We define C to be the set of non-anticipating functionals Z which are non-decreasing with $Z(0) \geq 0$, and we define C_x to be the set of $Z \in C$ such that $E_x R(Z) < \infty$ $(x \in \mathbb{R})$, where

$$R(Z) = \alpha \int_0^\infty e^{-\alpha t} Z(t) dt , \qquad Z \in C .$$

Then $P_x\{e^{-\alpha t} Z(t) \to 0 \text{ as } t \to \infty\} = 1$ for all $Z \in C_x$, and pathwise (Riemann-Stieltjes) integration by parts gives

$$P_x\{R(Z) = Z(0) + \int_0^\infty e^{-\alpha t} dZ(t)\} = 1 \qquad \text{for all } Z \in C_x .$$

We define $\mathcal{S}$ to be the set of (random) step functions in C having only finitely many jumps in any finite interval, and we take $\mathcal{S}_x$ to be the set of $Z \in \mathcal{S} \cap C_x$ such that $E_x R^*(Z) < \infty$ $(x \in \mathbb{R})$, where

$$R^*(Z) = 1_{\{Z(0) > 0\}} + \sum_{n=1}^\infty e^{-\alpha T_n(Z)} ,$$

where $T_n(Z)$ is the nth jump time of Z.

Proposition $E_x \sup\{e^{-\alpha t} Z(t) : t \geq 0\} < \infty$ if $x \in \mathbb{R}, Z \in C_x$.

<u>Proof</u>: Since Z is non-decreasing, we have

$$\int_t^\infty \alpha \ e^{-\alpha u} \ Z(u) du \geq \int_t^\infty \alpha \ e^{-\alpha u} \ Z(t) du = e^{-\alpha t} \ Z(t) \quad .$$

Taking the sup of each side over all $t \geq 0$, the left side becomes $R(Z)$, and the proposition follows from the fact that $E_x \ R(Z) < \infty$.

<u>Proposition 6</u>. Suppose $f : \mathbb{R}^+ \to \mathbb{R}$ is non-decreasing with $f(x) \leq a+bx$ for some $b > 0$, that $x \geq 0$, and that $Y \in C_x$ and $Z \in C_x$ satisfy

(5) $\qquad P_x\{X(t) + Y(t) - Z(t) \geq 0 \ \text{ for all } \ t \geq 0\} = 1$.

Then $E_x \ \sup\{e^{-\alpha t} \ f(X(t) + Y(t) - Z(t) : t \geq 0\} < \infty$.

<u>Proof</u>: Since f is non-decreasing and $Z(\cdot) \geq 0$, we have

(6) $\qquad e^{-\alpha t} \ f(0) \leq e^{-\alpha t} \ f(X(t) + Y(t) - Z(t)) \leq e^{-\alpha t} \ f(X(t) + Y(t))$

$$\leq e^{-\alpha t}[a + bX(t) + bY(t)]$$

almost surely (with respect to P_x) for all $t \geq 0$. The desired result then follows immediately from Propositions 1 and 5.

<u>Proposition 7</u>. Let f, x, Y and Z be as in Proposition 6. If $\{T_n, \ n \quad I\}$ are non-negative and $P_x\{T_n \to \infty\} = 1$ then

$$E_x \ e^{-\alpha T_n} \ f(X(T_n) + Y(T_n) - Z(T_n)) \to 0 \qquad \text{as} \quad n \to \infty \ .$$

<u>Proof</u>: From standard properties of Brownian Motion and the definition of C_x it follows that $e^{-\alpha t} X(t) \to 0$ and $e^{-\alpha t} Y(t) \to 0$ as $t \to \infty$ almost surely (with respect to P_x). Using the bound (6) and the fact that $T_n \to \infty$ almost surely, we then have

$$e^{\alpha T_n} \{X(T_n) + Y(T_n) - Z(T_n)\} \to 0 \qquad \text{as} \quad n \to \infty$$

almost surely. The desired result then follows, using Proposition 6 and dominated convergence.

3. The Control Problem with no Fixed Charges

For our first problem, we define an x-admissible policy $(x \geq 0)$ to be a pair of controls $Y, Z \in C_x$ satisfying (5). We say that an x-admissible policy (Y^*, Z^*) is x-optimal $(x \geq 0)$ if, for every other x-admissible policy (Y,Z),

$$E_x R(Y^*) - k E_x R(Z^*) \geq E_x R(Y) - k E_x R(Z) \quad .$$

We now construct a specific policy (Y^*, Z^*) which will eventually be shown x-optimal for all $x \geq 0$. Let $\beta = \mu/\sigma^2$ and $\gamma = (\beta^2 + 2\alpha/\sigma^2)^{1/2}$, and let S be the unique positive solution of

$$(\gamma-\beta) \ e^{(\gamma+\beta)S} + (\gamma+\beta) \ e^{-(\gamma-\beta)S} = 2k\gamma \quad .$$

It is easy to verify that the left side of (7) is continuous, convex and strictly increasing in S, with value 2γ at $S = 0$. Since $k > 1$, there is thus a unique positive solution. With $X(0) = x \geq 0$, we take $t_0 = T_0 = 0$ and $W_0(0) = 0$, and then for $n = 0, 1, \ldots$ we recursively define

$$X_{2n+1}(t) = W_{2n}(\tau_{2n}) + X(T_{2n}+t) - X(T_{2n}), \qquad t \geq 0 ,$$

$$Z_{2n+1}(t) = [\sup\{X_{2n+1}(u) : 0 \leq u \leq t\} - S]^+ , \qquad t \geq 0 ,$$

$$W_{2n+1}(t) = X_{2n+1}(t) - Z_{2n+1}(t) , \qquad t \geq 0 ,$$

$$\tau_{2n+1} = \inf\{t \geq 0 : W_{2n+1}(t) = 0\} \quad \text{and} \quad T_{2n+1} = T_{2n} + \tau_{2n+1} ,$$

$$X_{2n+2}(t) = X(T_{2n+1} + t) - X(T_{2n+1}) , \qquad t \geq 0 ,$$

$$Y_{2n+2}(t) = -\inf\{X_{2n+2}(u) : 0 \leq u \leq t\} , \qquad t \geq 0 ,$$

$$W_{2n+2}(t) = X_{2n+2}(t) + Y_{2n+2}(t) , \qquad t \geq 0 ,$$

$$\tau_{2n+2} = \inf\{t \geq 0 : W_{2n+2}(t) = S\} \quad \text{and} \quad T_{2n+2} = T_{2n+1} + \tau_{2n+2} .$$

Also, let $Y_1(t) = Z_2(t) = Y_3(t) = Z_4(t) = \cdots = 0$ for all $t \geq 0$. It is easy to show that

$$P_x\{0 = T_0 \leq T_1 \leq T_2 \leq \cdots \leq \infty\} = 1 \qquad \text{for all } x \geq 0 .$$

Then for $n \in I$ and $t \in [0, \tau_{n+1})$ we define

$$Y^*(T_n+t) = \sum_{i=1}^{n} Y_i(\tau_i) + Y_{n+1}(t) \; ,$$

$$Z^*(T_n+t) = \sum_{i=1}^{n} Z_i(\tau_i) + Z_{n+1}(t) \; ,$$

$$W^*(T_n+t) = W_{n+1}(t) \; ,$$

so that $W*(t) = X(t) + Y^*(t) - Z*(t) \geq 0$ for all $t \geq 0$. The initial time interval $[0, T_1]$ is a period of output control only. There is an initial output of size $Z^*(0) = Z_1(0) = [X(0)-S]^+$, and during the remainder of the period the cumulative output $Z^* = Z_1$ increases in the minimum amounts necessary to maintain $X-Z_1 \leq S$. The controlled process $W_1 = X-Z_1$ has state space $(-\infty, S]$ and $W_1(0) = [X(0) \wedge S]$, and it is known to behave as the Brownian Motion X modified by an upper (instantaneously) reflecting barrier at S. (We shall use this fact later without further comment.) The period ends when $W^* = W_1$ hits zero. Each subsequent interval of the form $[T_{2n}, T_{2n+1}]$ with $n \in I^+$ is similarly a period of output control only. The controlled process W^* starts in state S, the cumulative input Y^* remains constant, and the cumulative output $Z^{\cdot}$ increases in the minimum amounts necessary to maintain $W^* \quad X + Y^* - Z^{\cdot} \leq S$.

Each interval of the form $[T_{2n+1}, T_{2n+2}]$ is a period of input control only. The controlled process W^* starts in state zero and behaves as the Brownian Motion X modified by a lower reflecting barrier at zero. The cumulative output Z^* remains constant, and the cumulative input Y' increases in the minimum amounts necessary to maintain $W^* \geq 0$.

The period ends when W^* hits S. In total, W^* has state space $[0,S]$ and behaves as X modified by reflecting barriers at both boundaries.

Observe that (Y^*, Z^*) is an x-admissible policy for every starting state $x \geq 0$. We wish now to calculate

$$f(x) = E_x[R(Y^*) - kR(Z^*)], \qquad x \geq 0 .$$

As a first step, it is immediate from the construction that

(6) $\qquad f(x) = (x-S) + f(S) \qquad\qquad$ for $x \geq S$.

Assuming now that $X(0) \in [0,S]$, we define

$$f_1(x) = E_x[\alpha \int_0^{T_1} e^{-\alpha t} Z_1(t)dt + e^{-\alpha T_1} Z_1(T_1)] , \qquad 0 \leq x \leq S ,$$

$$A(x) = E_x e^{-\alpha T_1} , \qquad\qquad\qquad\qquad 0 \leq x \leq S .$$

Remembering that $Y^*(t) = 0$ if $0 \leq t \leq T_1$, it follows easily from our construction, the strong Markov property of X, and the definition of $R(\cdot)$ that

(7) $\qquad f(x) \quad f_1(x) + \alpha E_x \int_{T_1} e^{-\alpha t} [Z^*(t) - Z^*(T_1) - kY^*(t)]dt$

$\qquad\qquad\qquad f_1(x) + \alpha E_x e^{-\alpha T_1} f(W^*(T_1))$

$\qquad\qquad\qquad f_1(x) + \alpha A(x) f(0) , \qquad\qquad 0 \leq x \leq S .$

15

Now to solve for f_1 we recall that $W_1(t) = X(t) - Z_1(t)$ for $0 \leq t \leq T_1$ and $X(T_1) = Z_1(T_1)$ since $W_1(T_1) = 0$. Thus

$$(10) \qquad f_1(x) = \alpha E_x \int_0^{T_1} e^{-\alpha t}[X(t) - W_1(t)]dt + E_x e^{-\alpha T_1} X(T_1)$$

for $0 \leq x \leq S$. Defining $H(x) = E_x \int_0^{T_1} e^{-\alpha t} W_1(t)dt$ for $0 \leq x \leq S$, Proposition 3 and (10) give us

$$(11) \qquad f_1(x) = x + \mu[1 - A(x)]/\alpha - \alpha H(x) , \qquad 0 \leq x \leq S .$$

Recall that $T_1 = \inf\{t \geq 0 : W_1(t) = 0\}$ and that W_1 behaves as the Brownian Motion X with an upper reflecting barrier at S. Then standard results for the first passage times and potentials of Markov processes (in our case W_1) show that A and H satisfy the differential equations

$$(12) \qquad \mathcal{J} A(x) - \alpha A(x) = 0 \quad \text{and} \quad x + \mathcal{J} H(x) - \alpha H(x) = 0, \ 0 \leq x \leq S ,$$

with the boundary conditions

$$(13) \qquad A(0) = 1, \ H(0) = 0, \text{ and } A'(S) = H'(S) = 0 .$$

Combining (9) - (13) we find that f satisfies

$$(14) \qquad \mathcal{J}f(x) - \alpha f(x) = 0, \ 0 \leq x \leq S, \quad \text{and} \quad f'(S) = 1 .$$

14

Furthermore, there is a precisely symmetric argument, (defining a new sequence of stopping times T_n' such that the initial period $[0, T_n']$ is one of input control only when $0 < X(0) < S$) to show that the second boundary condition is

$$(15) \qquad\qquad\qquad f'(0) = k .$$

Using standard methods, the unique solution of (14) and (15) is found to be

$$(16) \qquad f(x) = \alpha e^{(\gamma - \beta)x} - b e^{-(\gamma + \beta)x} , \qquad 0 \leq x \leq S ,$$

where the constants a and b are chosen to satisfy the boundary conditions $f'(0) = k$ and $f'(S) = 1$. Elementary computations then give

$$(17) \qquad a = (e^{\beta S} - k e^{-\gamma S})/(\gamma - \beta)(e^{\gamma S} - e^{-\gamma S}) ,$$

$$(18) \qquad b = (k e^{\gamma S} - e^{\beta S})/(\gamma + \beta)(e^{\gamma S} - e^{-\gamma S}) .$$

Proposition 8. The function $f : \mathbb{R}^+ \to \mathbf{R}$ is concave, increasing and twice continuously differentiable with $f(S) = \mu/\alpha$. Furthermore, $\mathcal{L}f(x) - \alpha f(x) \leq 0$, $1 \leq f'(x) \leq k$ and $f(x) \leq (\mu/\alpha - S) + x$ for all $x \geq 0$.

Proof: From (3) and (14) it is immediate that f' exists and is continuous

on $\mathbb{R}^+$. For $0 < x < S$ we differentiate (16) twice to obtain

(19) $$f''(x) = a(\gamma-\beta)^2 e^{(\gamma-\beta)x} - b(\gamma+\beta)^2 e^{-(\gamma+\beta)x} .$$

Setting this expression equal to zero, multiplying through by $\exp(\beta x)$,

and substituting (17) and (18) for a and b, we find that f''(x) = 0

if and only if (7) holds with x in place of S. Thus f''(S) = 0 with

S chosen to satisfy (7). Clearly f''(x) = 0 for all $x \geq S$ by (8), so

f'' exists and is continuous. Since $|\beta| < \gamma$, it is clear from (18) that

b > 0. Since f''(S) = 0, it then follows from (19) that a > 0 and

hence that f'' is strictly increasing on [0,S]. Thus f''(x) < 0 for

$0 < x < S$, and it follows from (8) that f is strictly increasing and

concave on $\mathbb{R}^+$. Since f'(0) = k and f'(x) = 1 for $x \geq S$, this

implies that $1 \leq f'(x) \leq k$ for all $x \geq 0$. With f'(S) = 1 and

f''(S) = 0, we have $\mathcal{J}f(S) = \mu$. Since $\mathcal{J}f - \alpha f = 0$ on [0,S] this

gives us $f(S) = \mu/\alpha$. Then (8) yields $\mathcal{J}f(x) = \mu$ for $x \geq S$, implying

$$\mathcal{J}f(x) - \alpha f(x) = \mu - \alpha[\mu/\alpha + (x-S)] = -\alpha(x-S)$$

for $x \geq S$, so $\mathcal{J}f - \alpha f \leq 0$ on $\mathbb{R}^+$. Finally, $f(x) = (\mu/\alpha - S) + x$

for $x \geq S$ by (10), so the concavity of f gives $f(x) \leq (\mu/\alpha - S) + x$

for $x \geq 0$.

Remark. As the proof shows, f' is continuous regardless of how one

chooses S, but f'' is continuous if and only if S is chosen to

satisfy (7).

16

Theorem 1. If $x \geq 0$ and (Y,Z) is an x-admissible policy, then $E_x R(Z) - kE_x R(Y) \leq f(x)$. Thus the policy (Y^*, Z^*) constructed above is x-optimal for all $x \geq 0$.

Proof: Let $x \geq 0$ and (Y,Z) be fixed. When we speak of almost sure convergence, this refers to P_x. The idea of the proof is essentially to approximate Y and Z by (random) step functions. Given $\epsilon > 0$, we define an increasing sequence of stopping times T_n (hoping the reader will forgive this re-use of previous notation) by setting $T_0 \equiv 0$ and

$$T_{n+1} = \inf\{t \geq T_n : Y(t) \geq Y(T_n) + \epsilon \quad \text{or} \quad Z(t) \geq Z(T_n) + \epsilon \quad \text{or} \quad t = T_n + \epsilon\}$$

for $n \in I$. That each T_n is a stopping time follows immediately from the fact that Y and Z are nonanticipating functionals. Furthermore, $0 \leq T_1 < T_2 < \cdots$ because Y and Z are right continuous. Finally, $T_n \to \infty$ almost surely as $n \to \infty$ because both $R(Y)$ and $R(Z)$ are almost surely finite. Let

$$y_0 = Y(0) + \epsilon, \quad z_0 = Z(0), \quad q_0 = z_0 - ky_0 ,$$

$$A_0 = q_0 + f(X(0) + y_0 - z_0) - f(X(0)) ,$$

and then for $n \in I$

17

$$y_{n+1} = Y(T_{n+1}) - Y(T_n), \qquad z_{n+1} = Z(T_{n+1}) - Z(T_n),$$

$$Y_n = \sum_{i=0}^{n} y_i = Y(T_n) + \epsilon, \qquad Z_n = \sum_{i=0}^{n} z_i = Z(T_n),$$

$$D_n = Y_n - Z_n, \qquad q_n = z_n - ky_n,$$

$$A_{n+1} = e^{-\alpha T_{n+1}}[q_{n+1} + f(X(T_{n+1}) + D_{n+1}) - f(X(T_{n+1}) + D_n)],$$

$$B_n = e^{-\alpha T_{n+1}} f(X(T_{n+1}) + D_n) - e^{-\alpha T_n} f(X(T_n) + D_n).$$

With these definitions, we have

$$(20) \quad \sum_{i=0}^{n} A_i + \sum_{i=0}^{n-1} B_i = \sum_{i=0}^{n} e^{-\alpha T_i} q_i + e^{-\alpha T_n} f(X(T_n) + D_n) - f(X(0)).$$

From our construction and the fact that (Y,Z) is x-admissible it follows that $X(T_n) + D_n \geq \epsilon$ and $X(T_{n+1}) + D_n \geq 0$ almost surely for all $n \in I$. Since $1 \leq f'(x) \leq k$, it is then immediate that $A_n \leq 0$ almost surely for all $n \in I$. Furthermore, we can use Proposition 4 to show $E_x B_n \leq 0$ by making the following associations. Let f be as defined above for $x \geq 0$, and define $f(x)$ by (16) for $x \leq 0$. Then Proposition 8 shows that f satisfies the hypotheses of Proposition 4. Let $T^* = T_{n+1}$, $T_* = T_n$ and $Z = D_n = Y(T_n) - Z(T_n) + \epsilon$, so Z is measurable $\mathcal{F}_{T_*}$. Then Proposition 4 gives $E_x B_n \leq 0$, it following immediately from Propositions 6 and 8 that both expectations exist and are finite. Thus, taking the expectation of both sides in (20), we have

18

$$(21) \qquad E_x \sum_{i=0}^{n} e^{-\alpha T_i} q_i \le f(x) - E_x e^{-\alpha T_n} f(X(T_n) + D_n) .$$

Noting again that $D_n = Y(T_n) - Z(T_n) + \epsilon$, it follows from Propositions 7 and 8 that the second term on the right side of (21) goes to zero as $n \to \infty$, giving us

$$(22) \qquad E_x \sum_{i=0}^{\infty} e^{-\alpha T_i} q_i \le f(x) .$$

Let $Y^{\#}(t) = Y_n$ and $Z^{\#}(t) = Z_n$ for $T_n \le t < T_{n+1}$ and $n \in I$. From our construction it follows that $\langle Y^{\#}, Z^{\#} \rangle$ is an x-admissible policy with

$$Y(t) \le Y^{\#}(t) \le Y(t) + \epsilon \quad \text{and} \quad Z(t) - \epsilon \le Z^{\#}(t) \le Z(t)$$

for all $t \ge 0$, from which we have

$$(23) \qquad R(Z) - kR(Y) - (1+k)\epsilon \le R(Z^{\#}) - kR(Y^{\#})$$

$$= \sum_{i=0}^{\infty} e^{-\alpha T_i} q_i = R(Z) - kR(Y) .$$

Thus $E_x[R(Z) - kR(Y)] \le f(x) + (1+k)\epsilon$ by (22) and (23). Since $\epsilon > 0$ was chosen arbitrarily, the proof is complete.

4. The Control Problem with One Fixed Charge

For our second problem, we define an x-admissible policy $(x \geq 0)$ to be a pair of controls $Y \in \mathcal{D}_x$ and $Z \in C_x$ such that (5) holds. We say that an x-admissible policy (Y^*, Z^*) is x-optimal $(x \geq 0)$ if, for every other x-admissible policy (Y,Z),

$$E_x[R(Z^*) - kR(Y^*) - KR^*(Y^*)] \geq E_x[R(Z) - kR(Y) - KR^*(Y)] .$$

As in the previous section, let S be the unique positive solution of (7), and let $a > 0$ and $b > 0$ be defined in terms of S by (17) and (18) respectively. (Recall that we showed $a > 0$ in the proof of Proposition 8.) Now let s be the unique positive solution of

$$(24) \qquad a[1 - e^{-(\gamma-\beta)s}] + b[e^{(\gamma+\beta)} - 1] = K + ks .$$

Elementary calculations show that the left side of (24) is strictly convex and increasing with value zero at $s = 0$, and that its derivative increases without bound as s increases. Thus there is in fact a unique solution s.

With $X(0) \geq 0$, we set $\tau_0 = T_0 \equiv 0$ and then define $Z_1(t)$, $W_1(t)$ and $\tau_1 = T_1$ in terms of X and the positive constant $(s+S)$ exactly as they were defined in terms of X and the positive constant S in the previous section. Then for $n = 2, 3, \ldots$ let

$$X_n(t) = s + X(T_{n-1} + t) - X T_{n-1} , \qquad t \geq 0 ,$$

$$Z_n(t) = [\sup\{X_n(u) : 0 \leq u \leq t\} - s \cdot S]^+ , \qquad t \geq 0 ,$$

$$W_n(t) = X_n(t) - Z_n(t) , \qquad t \geq 0 ,$$

$$\tau_n = \inf\{t \geq 0 : W_n(t) = 0\} \quad \text{and} \quad T_n = T_{n-1} + \tau_n .$$

Again it is easy to show that

$$P_x\{0 = T_0 \leq T_1 < T_2 < \cdots \to \infty\} = 1 \qquad \text{for all} \quad x \geq 0 ,$$

and for $n \geq 1$ and $t \in [0, \tau_{n+1})$ we define

$$Y^*(t) = ns, \quad Z^*(t) = \sum_{i=1}^{n} Z_i(\tau_i) + Z_{n+1}(t) ,$$

and $W^*(t) = W_{n+1}(t)$, so that $W^*(t) = X(t) + Y^*(t) - Z^*(t)$ for all $t \geq 0$. The behavior of the controlled process $W^* = W_1$ during the initial interval $[0, T_1)$ is as described in Section 3 except that now the upper reflecting barrier is at $(s+S)$. Each subsequent time interval $[T_n, T_{n+1})$ begins with a jump of s in the cumulative input Y^*, this moving the controlled process W^* from state zero to state s. During the remainder of the period, Y^* remains constant and the cumulative output Z^* increases in the minimum amounts necessary to maintain $W^* \leq (s+S)$. The period ends when W^* hits zero and jumps upward by s again.

Using arguments very much like those employed in the previous section, it is straight-forward to show that the function

$$g(x) = E_x[R(Z^*) - kR(Y^*) - KR^*(Y^*)], \quad x \geq 0,$$

satisfies $\mathcal{L}g(x) - \alpha g(x) = 0$ for $0 \leq x \leq s+S$, with $g'(s+S) = 1$, $g(0) = g(s) - K - ks$, and $g(x) = g(s+S) + (x-S-s)$ for $x \geq s+S$. Again the unique solution has the form

$$(25) \qquad g(x) = a_0 \, e^{(\gamma-\beta)x} - b_0 \, e^{-(\gamma+\beta)x} \qquad \text{for } 0 \leq x \leq s+S,$$

where the constants a_0 and b_0 are selected to satisfy the boundary conditions $g'(s+S) = 1$ and $g(0) = g(s) - K - ks$. The reader may easily verify that, with s chosen to satisfy (24), the selection

$$(26) \qquad a_0 = a \, e^{-(\gamma-\beta)s} \qquad \text{and} \qquad b_0 = b \, e^{(\gamma+\beta)s}$$

meets the boundary condition $g(0) = g(s) - K - ks$. We then further have (defining f by (16))

$$(27) \qquad g(s+x) = a \, e^{(\gamma-\beta)x} - b \, e^{-(\gamma+\beta)x} = f(x) \qquad \text{for } 0 \leq x \leq S.$$

In Section 3 we showed that $f'(S) = 1$, so our second boundary condition $g'(s+S) = 1$ is also satisfied, and the complete solution for g is given by (25), (26) and $g(s+S+x) = g(s+S) + x$ for $x \geq 0$.

Proposition 9. The function $g : \mathbb{R}^+ \to \mathbb{R}$ is concave, increasing and twice continuously differentiable with $g(s+S) = \mu/\alpha$. Furthermore, $\beta g(x) - \alpha g(x) \leq 0$, $g'(x) \geq 1$ and $g(x) \leq (\mu/\alpha - s - S) + x$ for all $x \geq 0$. Finally, $g(x+y) - g(x) \leq K + ky$ for all $x \geq 0$ and $y \geq 0$.

Proof: Clearly $a_0 \geq 0$ and $b_0 > 0$, and differentiating (25) twice gives

$$(26) \qquad g''(x) = a_0 (\gamma-\beta)^2 \, e^{(\gamma-\beta)x} - b_0 (\gamma+\beta)^2 \, e^{-(\gamma+\beta)x}, \qquad 0 \leq x \leq s+S \, ,$$

so g'' is strictly increasing on $[0, s+S]$. We showed in the proof of Proposition 8 that $f(S) = \mu/\alpha$, $f'(S) = 1$ and $f''(S) = 0$, so (27) gives us $g(s+S) = \mu/\alpha$, $g'(s+S) = 1$ and $g''(s+S) = 0$. It then follows as in the proof of Proposition 8 that g is concave, increasing and twice differentiable with $\beta f - \alpha f \leq 0$. Since $g'(x) = 1$ for $x \geq s+S$, the concavity implies $g'(x) \geq 1$ for all $x \geq 0$. Also, we showed in Section 3 that $f'(0) = k$, so (27) gives $g'(s) = k$. From the concavity of g we then have

$$(27) \qquad g(y) - g(0) - ky \leq g(s) - g(0) - ks = K \qquad \text{for all } y \geq 0 \, .$$

Finally, the concavity of g implies that $g(x+y) - g(x)$ is a non-increasing function of x for all $y \geq 0$. Combining this with (29) proves the last statement of the proposition.

Theorem 2. If $x \geq 0$ and (Y,Z) is an x-admissible policy, then
$E_x[R(Z) - kR(Y) - KR^*(Y)] \leq f(x)$. Thus the policy (Y^*, Z^*) constructed
above is x-optimal for all $x \geq 0$.

Proof: The proof is very similar to that of Proposition 8, and we shall
only sketch it. We define the sequence of stopping times T_n (again
apologizing for the notation) by $T_0 = 0$ and

$$T_{n+1} = \inf\{t \geq T_n : Y(t) > Y(T_n) \quad \text{or} \quad Z(t) \geq Z(T_n) + \epsilon$$

$$\text{or} \quad t = T_n + \epsilon\}$$

for $n \in I$. We proceed exactly as in the proof of Proposition 8
except that f is replaced by g throughout, we take $y_0 = Y(0)$,
and we define

$$q_n = z_n - ky_n - K \, 1_{\{y_n > 0\}} \quad , \quad n \in I$$

To show that $A_n \leq 0$ almost surely with this change in the definition of
q_n, we use the fact that $g' \geq 1$ and $g(x+y) - g(x) \leq K + ky$ by
Proposition 9. That $E_x B_n \leq 0$ for all n follows exactly as before.
Defining $Y^{\#}(t) = Y_n$ and $Z^{\#}(t) = Z_n$ for $T_n \leq t < T_{n+1}$ and $n \in I$,
we arrive at

$$E_x[R(Z^{\#}) - kR(Y^{\#}) - KR^*(Z^{\#})] = E_x \sum_{i=0}^{\infty} e^{-\alpha T_i} q_i \leq f(x) \quad .$$

But $Y^{\#}(t) = Y(t)$ and $Z(t) - \epsilon \leq Z^{\#}(t) \leq Z(t)$ for all $t \geq 0$, so the
desired result follows directly.

24

. Applications

Consider an inventory and production system involving a single
type of item (product), and assume that the cumulative excess production
of the item can be reasonably represented by the Brownian Motion
$X = \{X(t), t \geq 0\}$. We have in mind a situation where there is a non-
decreasing cumulative production process $P = \{P(t), t \geq 0\}$ and a non-
decreasing cumulative demand process $D = \{D(t), t \geq 0\}$ such that $P-D$
can be approximated by X. We interpret P as the production from
regular operations and assume that additional instantaneous increases
in the stock level can be accomplished by some irregular means (such
as overtime production or ordering from an outside vendor) at a cost of
$k > 0$ dollars per item. We interpret D as the demand from regular
customers and assume that unlimited additional quantities of (excess)
stock can be sold by irregular means (such as sale to a scavenger as
scrap) at a price of c dollars per item, where $0 < c < k$. Let
$Y = \{Y(t), t \geq 0\}$ and $Z = \{Z(t), t \geq 0\}$ denote the cumulative
irregular production and cumulative irregular sales respectively.
Assuming that all regular demand must be met instantaneously (no
backlogging), the controls Y and Z must be chosen so that the stock
level $W(t) = X(t) + Y(t) - Z(t)$ is non-negative for all $t \geq 0$.
Assuming that inventory holding costs are continuously incurred at
rate $h W(t)$, and that future costs and revenues are continuously
discounted at interest rate $\gamma > 0$, we wish to choose Y and Z so
as to maximize

$$E\{cR(z) - kR(Y) - \int_0^\infty e^{-\alpha t} h(W(t))dt\} .$$

If $h(x) = hx$ for all $x \geq 0$, then the last term inside the brackets
is just

$$h \int_0^\infty e^{-\alpha t}[X(t) + Y(t) - Z(t)]dt = (h/\alpha) [R(Y) - R(Z)] + h \int_0^\infty e^{-\alpha t} X(t)dt.$$

Observing that the last term is uncontrollable (depends on neither Y
nor Z), we then see that the total objective is to maximize $(c + h/\alpha) ER(Z)$
$- (k + h/\alpha) ER(Y)$. Assuming without loss of generality that $c + h/\alpha = 1$,
this is precisely the problem formulated in Section 3. If the irregular
production process Z is a step function, and if a set-up cost of $K > 0$
is incurred each time that a jump in the irregular production occurs,
then we similarly obtain the problem formulated in Section 4.

A closely related diffusion model of optimal inventory control
has been advanced by Bather (1966). In our notation, Bather assumes
$\mu < 0$ and $K > 0$, and he considers a linear holding cost function $h(\cdot)$.
There is no provision for sale of excess inventory in his model
$(Z(\cdot) = 0)$. The stock level is permitted to go negative (backlogging
is permitted), but linear shortage costs are incurred when this happens.
Attention is restricted to a simple class of input step functions Y
which jump by a fixed amount (S-s) whenever the inventory on hand
decreases to a fixed level s, the objective being to minimize <u>average</u>
cost per unit time over an infinite planning horizon. Other diffusion
formulations for problems of optimal control of dams, inventories and
storage systems are given by Bather (1968), Faddy (1974a,b), Whitt (1973a,b),

Puterman (1977), and Priska (1978). In all of these papers, attention
is restricted to (non-randomized and Markov) stationary policies, and
in all but the last there is a further restriction to stationary policies
having a particular structure.

As a second application, we consider the stochastic cash management
problem, discrete-time versions of which have been studied by Eppen and
Fama (1969), Girgis (1968), and Neave (1970). Imagine a firm which
maintains a cash fund, into which a certain amount of income or revenue
is automatically channeled and out of which operating disbursements are
made. We assume that the resulting fluctuations in the content of the
fund can be adequately represented by the Brownian Motion X. Additional
instantaneous increases in the content of the fund can be accomplished
by converting securities into cash, but there is a transaction cost of
$k > 0$ dollars incurred for each dollar of securities so converted.
Also, cash from the fund can be converted into securities at a trans-
action cost of $c > 0$ dollars per dollar so converted. Finally an
opportunity loss of $h > 0$ dollars per unit time is suffered for each
dollar that is held within the fund. Denoting by $Y(t)$ the cumulative
conversion of securities to cash up to time t, and by $Z(t)$ the cumulative
conversion cash to securities, we assume that the content of the fund
must be kept non-negative. Then the problem is to find an admissible
policy (Y,Z) that maximizes $(h/\alpha - c) ER(Z) - (h/\alpha - k) ER(Y)$, where
the linear opportunity cost (holding cost) has been converted as in the
previous example. If $c < h/\alpha$ and $c < k$, then we may assume without
loss of generality that $h/\alpha - c = 1$, and we have precisely the problem
formulated in Section 5. If the conversion of securities to cash entails

both a fixed charge $K > 0$ and the proportional charge k, then we get the problem formulated in Section 4.

Constantinides (1976) has examined both of these cash management problems with the objective of minimizing _average_ _cost_ (rather than expected discounted cost) over an infinite planning horizon, and he has further considered the case where both types of conversion entail both fixed and proportional transaction costs. (See Section 6.) His results are very similar to ours, but the methodology employed is quite different, and we do not fully understand his proofs.

For both inventory control problems and stochastic cash management problems, one finds that the diffusion formulations discussed above are much more tractable than more traditional (usually discrete review) models. For our particular problems, we have shown that the optimal policy (from a very broad class of potential policies) has an extremely simple structure, and the assumption of an underlying Brownian Motion further permits explicit determination of the relevant critical numbers. With traditional formulations, even the structural results may fail, and the computation of optimal policies is typically a complicated matter. See Girgis (1968) for a demonstration of this in the case of cash management.

On the other side of the issue, it simply may not be reasonable to represent the underlying (net demand or net production) process by a Brownian motion. Bather (1966) suggested that a non-decreasing demand process be approximated by Brownian Motion with positive drift, but as Whitt (1973a) has pointed out, this is not a circumstance where one expects a good approximation. In each of our applications, we have

emphasized problems where the underlying process X represents the

difference of two non-decreasing processes, and in this circumstance

various theorems on the (weak) convergence of stochastic processes may

be invoked (with further assumptions on the parameters of the relevant

processes) to justify the Brownian approximation. See, for example,

Harrison (1975). Even when this can be done, there remains the problem

of justifying one's diffusion _optimization_ problem as a reasonable

approximation to the original optimization problem. If one restricts

attention to policies which are explicit functionals (and continuous in

the appropriate function space topology) of the underlying process, we

believe that this might be a manageable task in the rather restrictive

setting of our model, but the issue will not be pursued further here.

6. __Concluding Remarks__

If for each $t \geq 0$ we define $\mathcal{B}_t$ to be the sub-σ-algebra

generated by $\{X(u), 0 \leq u \leq t\}$, then it follows from the stationary,

independent increments property of X that $X(t+u) - X(t)$ is independent

of $\mathcal{B}_t$ for all $t \geq 0$ and $u \geq 0$. Thus we may take $\mathcal{F}_t = \mathcal{B}_t$ in our

formulation (Section 2), and in general we must have $\mathcal{B}_t \subset \mathcal{F}_t$ for all

$t \geq 0$. If a control policy (Y, Z) has the property that $Y(t)$ and

$Z(t)$ are measurable $\mathcal{B}_t$ for all $t \geq 0$, we shall call it a _non-randomized_

policy. Given a non-randomized policy (Y, Z), let $\mathcal{N}_t$ be the σ-algebra

in $\mathcal{B}_t$ generated by $X(t) + Y(t) - Z(t)$ for $t \geq 0$. We shall say that

(Y, Z) is a _Markov policy_ if $Y(t+u)$ and $Z(t+u)$ are measurable $\mathcal{N}_t$

for all $t \geq 0$ and $u \geq 0$. Finally, we shall say that a Markov policy (Y,Z) is _stationary_ if the conditional distribution of $(Y(t+u) - Y(t),$ $Z(t+u) - Z(t); u \geq 0)$ given $\mathcal{N}_t$ depends only on $X(t) + Y(t) - Z(t)$ and not on t $(t \geq 0)$. All of this terminology conforms with the standard usage in (discrete time) dynamic programming, appropriately adapted to our setting. Roughly speaking, a policy (Y,Z) is non-randomized if the controls applied up to time t depend only on the history of the underlying process X up to time t and not on any other "irrelevant" information contained in $\mathcal{J}_t$. It is Markov if the controls to be applied after time t depend on the history up to time t only through the current "state of the system" $W(t) = X(t) + Y(t) - Z(t)$, and it is stationary if this dependence on $W(t)$ does not involve t. As the reader may easily verify, the policies that we have shown to be optimal for our two problems are both (non-randomized and Markov) stationary policies.

A natural successor for the two problems considered in this paper is one where both the input control Y and the output control Z must be step functions, and there are (different) fixed costs associated with both input jumps and output jumps. Similar (and much more complex) problems of pure impulse (jump) control have been considered by Bensoussan and Lions (1973, 1975) and by Richard (1976), but the method of proof used here can also be extended to the case of two fixed charges. As we shall demonstrate in a future paper, there exists an optimal solution for this problem which involves only three critical numbers, but the computation of those critical numbers is quite complicated. If we allow one (respectively, both) of the fixed

charges to approach zero, we find that the optimal controls approach
those displayed in Section 4 (respectively, Section 3) almost surely.
Thus, roughly speaking, each of the problems treated here can be obtained
as the limit of problems involving two fixed charges.

Our problem with no fixed charges (Section 3) can also be approxi-
mated by a formulation of the type considered by Mandl (1968) and Pliska
(1973). Suppose that the non-decreasing controls Y and Z are both
required to be absolutely continuous and non-decreasing with a density
bounded by $c > 0$. We cannot then require that $W(t) = X(t)+Y(t)-Z(t)$
remain positive, but we suppose that a large penalty cost of M (dollars
per unit time) is continuously incurred so long as $W(t) < 0$. It can
then be shown that there are critical numbers a and b with $0 < a < b < \infty$
such that one optimal policy is the following. When $W(t) < a$, the con-
troller increases Y at the maximum permissible rate c, when $W(t) > b$
he increases Z at rate c, and when $a \leq W(t) \leq b$ he does nothing.
If we let $c \to \infty$, we find that $a \to 0$, $b \to S$ and the optimal controls
converge almost surely to those displayed in Section 3.

Having indicated in the last two paragraphs that our problems can
be approximated in either of two ways, we emphasize that either type of
approximate formulation is harder to solve than the problems as we have
stated them. Also, we repeat that the optimal controls displayed in
Sections 3 and 4 are neither absolutely continuous nor step functions.
We are not aware of any previous formulation of a stochastic control
problem which permits controls of the type that we have found to be
optimal.

Acknowledgement

References

[1] Bather, J.A. (1966), A Continuous time Inventory Model, *J. Appl. Prob.*, 3, 538-549.

[2] Bather, J.A. (1968), A Diffusion Model for the Control of a Dam, *J. Appl. Prob.*, 5, 55-71.

[3] Bensoussan, A., and J.-L. Lions (1973), Nouvelle Formulation de Problemes de Controle implusionnel et Applications, *Comptes Rendus*, Academie des Sciences, Paris, Vol. 276, Series A, 1189-1192.

[4] Bensoussan, A., and J.-L. Lions (1975), Nouvelle Methodes in Controle Impulsionnel, *J. of Appl. Math. and Opt.*, 1,

[5] Breiman, L. (1968), *Probability*, Addison-Wesley, Reading, Mass.

[6] Constantinides, G.M. (1974), Stochastic Cash Management with Fixed and Proportional Transaction Costs, *Mgt. Sci.*, 22, 1320-1331.

[7] Eppen, G.D., and E.F. Fama (1968), Solutions for Cash Balance and Dynamic Portfolio Problems, *J. of Business*, 41,

[8] Faddy, M.J. (1974a), Optimal Control of Finite Dams: Discrete (2-Stage) Output Procedure, *J. Appl. Prob.*, 11, 111-121.

[9] Faddy, M.J. (1974b), Optimal Control of Finite Dams, *Adv. Appl. Prob.*, 6, 689-710.

[10] Girgis, N.M. (1968), Optimal Cash Balance Levels, *Mgt. Sci.*, 15,

[11] Harrision, J.M. (1975), Risk Processes with Compounding Assets, Technical Report, Dept. of Ops. Rsch., Stanford Univ. (to appear in Sto. Pro. and Appl.).

[12] Mandl, P. (1968), Analytical Treatment of One-Dimensional Markov Processes, Springer-Verlag, New York.

[13] Neave, E. H. (1970), The Stochastic Cash Balance Problem with Fixed Costs for Increases and Decreases, Mgt. Sci., 16,

[14] Pliska, S.R. (1973), Single Person Controlled Diffusions with Discounted Costs, J. Opt. Th. Appl., 12, 248-255.

[15] Pliska, S.R. (1975), A Diffusion Model for the Optimal Operation of a Reservoir System, J. Appl. Prob., 12, 859-863.

[16] Puterman, M.L. (1975), A Diffusion Process Model for a Storage System, Logistics, ___, 143-159.

[17] Richard, S.F. (1975), Optimal Impulse Control of a Diffusion Process with Both Fixed and Proportional Costs of Control, Technical Report, Grad. Sch. of Ind. Adm., Carnegie-Mellon Univ. (to appear in SIAM J. of Control).

[18] Whitt (1973a), Diffusion Models for Inventory and Production Systems, Technical Report, Dept. of Administrative Sciences, Yale Univ.

[19] Whitt (1973b), The Diffusion Inventory Model with Discounting, Technical Report, Dept. of Administration Sciences, Yale Univ.

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER TR - 69	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE *(and Subtitle)* OPTIMAL CONTROL OF A BROWNIAN STORAGE SYSTEM.		5. TYPE OF REPORT & PERIOD COVERED Technical Report
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) J. Michael Harrison Allison J. Taylor		8. CONTRACT OR GRANT NUMBER(s) N00014-76-C-0418, NSF-E -75-14847
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Operations Research Stanford University Stanford, CA 94305		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR-047-061
11. CONTROLLING OFFICE NAME AND ADDRESS Operations Research Program, Code 434 Office of Naval Research Arlington, Virginia 22217		12. REPORT DATE Sept 76
		13. NUMBER OF PAGES 33
14. MONITORING AGENCY NAME & ADDRESS*(if different from Controlling Office)*	37p.	15. SECURITY CLASS. *(of this report)* Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT *(of this Report)*

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT *(of the abstract entered in Block 20, if different from Report)*

19. KEY WORDS *(Continue on reverse side if necessary and identify by block number)*

Stochastic Control	Diffusion Processes
Optimal Control of Diffusions	Stochastic Cash Management
Inventory and Production Control	Reflecting Brownian Motion

20. ABSTRACT *(Continue on reverse side if necessary and identify by block number)*

See reverse side

"Optimal Control of a Brownian Storage System", by J. Michael Harrision
and Allison J. Taylor

$> \text{or} =$

Abstract

Consider a storage system (such as an inventory or bank account) whose
content fluctuates as a Brownian Motion $X = \{X(t), t \geq 0\}$ in the absence of
any control. Let $Y = \{Y(t), t \geq 0\}$ and $Z = \{Z(t), t \geq 0\}$ by non-
decreasing, non-anticipating functionals representing the cumulative input
to the system and cumulative output from the system respectively. The
problem is to choose Y and Z so as to maximize expected discounted
reward subject to the requirement that $X(t) + Y(t) - Z(t) \geq 0$ for all
$t \geq 0$ almost surely. In our first formulation, we assume that a reward of
one dollar is received for every unit of output, while a cost of $k > 1$
dollars is incurred for every unit of input. We explicitly compute an
optimal policy involving a single critical number. In our second formula-
tion, the cumulative input Y is required to be a step function, and an
additional cost of $K > 0$ dollars is incurred each time that an input
jump occurs. We explicitly compute an optimal policy involving two
critical numbers. Applications to inventory/production control and
stochastic cash management are discussed.